2005

Active Multimodal Presentations, an overview

Adel Elsayed
University of Bolton, a.elsayed@bolton.ac.uk

Active Multimodal Presentations, an overview

Adel Elsayed
University of Bolton
a.elsayed@bolton.ac.uk

Abstract

This workshop is concerned with a new approach to constructing effective, stand alone, multimodal presentational content, referred to as Active Multimodal Presentation (AMP). This article provides a brief overview of this new type of content, outlines the basic differences between modes of multimodal presentations in general and the tools required for developing them. AMP, as a special type of computer-based multimodal presentation is then discussed, to show that while a bespoke application will make use of the full potential of AMP technology, simpler solutions, based on tools that are readily available, could provide immediate access to such effective means of producing presentations.

1. Introduction

Active multimodal presentations (AMPs) are stand-alone, free running, presentations that *optimize the utilization* of the principal perceptual modalities of the audience, i.e. the auditory and visual modes. This should enhance the effectiveness of presentations, which is particularly useful for achieving their educational objectives. The optimization in this context can be realized in two ways, *mode assignment* and *integration of modalities*. In the former case, mode assignment, the message content is distributed amongst modalities: a *verbal component* that addresses the auditory mode, and a *pictorial component* that addresses the visual mode. In addition, *gestures* can be considered as a third modality that supports and enhances the semantic content of the message. Gestures also provide an instrument for integrating message modalities, which is the second means of optimization.

AMPs, therefore, have three main components: an auditory component, representing *speech narrative*, a visual component, representing *visual objects* and a gestural component representing the *integrating object*, which jells together the other two components into a *coherently integrated multimodal* (CIM) environment [1]. In human communication, the natural integrating object is the *hand*. We use our hands, not only to externalize our internal representations, by scripting, sketching etc, but also to add a *gesture* component to our communication act that directs attention, provides illustration, emphasizes the contours of the narrative, as well as expressing affect and empathy. The role of gestures, therefore, is not confined to the integration of message modalities, but extends to a *social level*, which integrates together the presenter and the audience into a socially coherent communication environment.

One of our research themes is concerned with the analysis and understanding of the nature and role of these gestural attributes in various types of multimodal presentations [2]. One of these is the computer-based presentation mode that we refer to as an AMP, the version introduced in this workshop. The other versions include board-based presentations and table-based presentations, captured through video cameras. AMPs are usually captured through a *screen capturing* software utility, sometimes referred to as a *Screen Cam* or by using a dedicated software presentation tool, specifically developed for that purpose.

As a preface to the workshop activities, the next section outlines the technology used for developing general multimodal presentations, from which the requirements for developing AMP content will be identified. The development of AMPs will then be explored in section 3. Section 4 provides a summary of the paper.

2. Developing Multimodal presentations

Reflection on the differences between the above three methods of capturing presentations, video, screen cam and bespoke software would reveal the motivation for developing AMP technology. These differences relate to three main factors, one is concerned with *resource* issues; the second is concerned with issues of tractability to *machine processing*; and the third relates to *production flexibility*, an important factor that influences cost, and quality.

As for the resource aspect, which relates to communication resources, video is known to be a

Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05)
0-7695-2338-2/05 $20.00 © 2005 **IEEE**

highly demanding medium for streaming bandwidth. Hence, there would be a natural interest in developing a medium that is as effective as video, but not as demanding on bandwidth requirements. The solution lies in AMP, for reasons indicated hereafter.

In terms of tractability to machine processing, video is a *scanning-based* acquisition technology, which does not lend itself to readily identifying the individual components that make up the captured scene. In fact, this has become the objective of another research area, known as computer vision, which aims at identifying and manipulating individual components. In addition, there is some research interest directed towards 'component-based encoding' for interactive television. Using these new technologies, however, would entail adding unnecessary complications and costs to the process when it is possible to *compose* the scene out of the constituting components at the point of production.

Video, therefore, is not the right technology for a *composition-based production* approach. AMP on the other hand is a *synthetic* medium, hence naturally suited for this kind of *flexible production* approach; in AMP, every modality is independently acquired and, hence, is conveniently encoded for full accessibility for machine processing. The independent acquisition of AMP modalities provides a sound basis for developing content acquisition software that fits natural human requirements for the delivery of presentations. This feature ensures that content acquisition could be realized *efficiently*, hence minimizing production costs, yet very *effective* in its communication attributes. In addition, the component-based approach would allow every modality of content to be independently optimized for *streaming bandwidth*, hence reducing bandwidth requirements to a minimum.

The pedagogical aspects of multimodal presentations are discussed in another article [2].

3. Development of AMP content

Part of our research, at the APT Lab, aims at developing software that addresses the problems of AMP acquisition, streaming and delivery, while another part is concerned with investigating the characteristics and attributes of AMP components e.g. gestures' roles and attributes, characteristics of AMP discourse, semantic processing of AMP content etc. As our target software is still under development, we have identified compromised solutions for the creation of AMP content using generic multimedia tools e.g. Macromedia's Director.

In spite of its limitations, Director provides a convenient platform for prototyping AMPs e.g. for empirical proof-of-concept investigations. The generic nature of Director, however, meant that its IDE (integrated development environment) has become so complex that a specialist would usually be required to use it. A simpler alternative can be found in screen capturing utilities e.g. ScreenCam, HyperCam, Camtesia etc. This is a straight forward option to use that does not need any special training, yet is capable of producing content that makes use of gestures in communication and, at the same time, is moderate in its demand for bandwidth. Its drawback, however, is the lack of component accessibility for machine processing. As a research objective, this point should not hinder the possibility of using such compromised solution for producing a useful form of AMP content, though it may be lacking in some advanced features.

In fact, such simple solutions, referred to as *telepointers* in the literature, has been found to be very effective in directing attention, creating real presence and improve learning performance [3]. The use of screen capturing utilities to capture *on-screen pointing* when developing presentational content for learning is explored in more detail in another article in the workshop.

4. Summary

This article has provided a brief overview of Active Multimodal presentations, indicating some basic differences between modes of multimodal presentations and the tools required for developing them. It then elaborated on the development of AMP content, to show that while a bespoke application will make use of the full potential of AMP technology, simpler solutions could provide immediate access to such effective means of presentations. These techniques will be explored in the workshop activities.

References

[1]Elsayed M.A.N. "A New Communication Paradigm: the coherently integrated multimodal medium for narratives", Technical Report, LSI, Leeds, UK, May 2001.

[2] Hartley R., Elsayed A., Pesheva M. "The Anatomy of a presentation: Syntactic and Semantic Analysis of an Active Multimodal Presentation" 5th IEEE International Conference on Advanced Learning Technologies (ICALT 2005), Kaohsiung, Taiwan July 5-8, 2005.

[3] Adams J, Rogers B, Hayne S, Mark G, Nash J, Leifer L. "The effect of a telepointer on student performance and preference" Computers & Education 44 (2005) 35-51.